

Université de Montréal

Rapport de recherche

La pertinence de l'évaluation hors-échantillon  
du modèle SETAR(3) appliqué aux taux de change réels

Rédigé par :  
Sergiu Luca

Dirigé par :  
Prof. Marine Carrasco

Département de sciences économiques  
Faculté des arts et des sciences

Le 7 février 2013

# Table des matières

<b>I.</b>	<b>L'INTRODUCTION .....</b>	<b>2</b>
<b>II.</b>	<b>LE RÉSUMÉ DE LA LITTÉRATURE.....</b>	<b>2</b>
<b>III.</b>	<b>LA DESCRIPTION DES DONNÉES.....</b>	<b>8</b>
<b>IV.</b>	<b>L'ESTIMATION ET LA PRÉVISION DU MODÈLE SETAR À TROIS RÉGIMES .....</b>	<b>16</b>
<b>1.</b>	<b>L'estimation .....</b>	<b>16</b>
<b>2.</b>	<b>La prévision.....</b>	<b>23</b>
<b>i.</b>	<b>La puissance à niveau corrigé (PNC).....</b>	<b>25</b>
<b>V.</b>	<b>LA CONCLUSION.....</b>	<b>28</b>

## I. L'INTRODUCTION

Lorsque nous essayons de modéliser l'évolution d'une variable économique (ex. taux de change) dans le temps, et nous avons le choix entre quelques modèles, il faut que nous soyons capables d'en choisir le meilleur. C'est-à-dire que nous cherchons le modèle qui décrit mieux la dynamique de la variable visée. À cet effet, nous pouvons comparer les prévisions hors-échantillon provenant de différents modèles. Cependant, si les prévisions obtenues à partir d'un de ces modèles sont en moyenne plus proches des vraies valeurs que les prévisions des autres modèles, signifie-t-il que ce modèle est le meilleur parmi les modèles considérés?

Dans ce travail nous répondons à cette question pour un modèle autorégressif au seuil auto-stimulant (SETAR) à trois régimes. La variable macroéconomique visée est le taux de change réel, et le test utilisé pour comparer des prévisions concurrentes est Diebold et Mariano (1995)(DM). Notamment, nous comparons les prévisions ponctuelles à une période provenant d'un modèle SETAR à trois régimes avec celles provenant d'une marche aléatoire. Ainsi, nous trouvons la puissance du test DM et nous la comparons avec la puissance de test de racine-unitaire proposé par Bec, Ben Salem et Carrasco (2004).

## II. LE RÉSUMÉ DE LA LITTÉRATURE

L'article par Diebold (2012) joue un rôle central pour ce travail. Dans cet article l'auteur affirme clairement que l'objectif principal du test DM est la comparaison des prévisions plutôt que des modèles. Dans ce contexte, les prévisions doivent être traitées comme primitives. C'est-à-dire que la source de prévisions n'a pas de pertinence. En fait, la source de prévisions n'est pas nécessairement un modèle; il peut s'agir, entre autres, de l'information obtenue à partir du marché de capitaux ou des enquêtes. Il faut toujours garder à l'esprit que ce test a été conçu pour établir si la différence entre les erreurs de prévision échantillonnelles de deux séries de prévisions est statistiquement significative. Notamment, le test DM établit si la différence est particulière à une réalisation d'échantillon ou si la différence est en fait présente dans la population. Ainsi, le test DM performe bien si nous comparons exclusivement des prévisions, mais il n'est pas recommandé pour l'évaluation de modèles.

Ensuite, Diebold (2012) fait ressortir que le test DM, appliqué adéquatement, est axé sur une seule hypothèse, soit la stationnarité de second ordre de la série de la différence des fonctions de perte. Dans la deuxième partie de l'article, il y a une discussion détaillée dont l'objectif est de réitérer les avantages

d'une évaluation dans-échantillon par rapport à une évaluation (pseudo-)hors-échantillon quand nous comparons les modèles en soi. Ce message est renforcé par le fait que la taille d'échantillon typique en macroéconomie est généralement loin de l'infini de la théorie asymptotique.

L'article par Bec, Ben Salem et Carrasco (2004) est notre guide dans le choix du modèle. Les auteurs trouvent qu'un modèle SETAR à trois régimes décrit bien la dynamique de taux de change réel. La raison est liée à la présence d'une région où la parité de pouvoir d'achat (PPA) ne s'applique pas. Autrement dit, dans cette région le taux de change réel ne tend vers la valeur imposée par la PPA. Dans la théorie, selon la PPA, sous certaines conditions les niveaux des prix exprimés dans une devise commune doivent être égaux. Donc, nous nous attendons à ce qu'une déviation crée la possibilité d'arbitrage qui retournera le taux de change réel au niveau imposé par la PPA. Quand même, la PPA ne tient pas toujours. Par exemple, si le bénéfice provenant de l'arbitrage est plus petit que les coûts de transaction, le mécanisme de la PPA ne sera pas activé à travers l'exploitation de la possibilité d'arbitrage. Par conséquent, il y aura une région où le taux de change réel varie sans tendance vers sa valeur de la PPA. Dans l'article de Bec, Ben Salem et Carrasco (2004) nous trouvons d'autres raisons proposées dans la communauté académique qui expliquent l'échec de la PPA.

Avant d'appliquer le modèle SETAR à trois régimes aux taux de change réel les auteurs préparent le terrain en abordant des aspects théoriques. Dans le modèle SETAR à trois régimes, le régime de milieu représente la région où la PPA ne s'applique pas. Ainsi il est possible pour un modèle SETAR à trois régimes d'être stationnaire et mélangeant même s'il y a une racine unitaire dans le régime de milieu. Ceci est démontré en premier lieu par Bec, Ben Salem et Carrasco (2004). L'article propose un test développé spécifiquement pour tester l'hypothèse nulle de la racine unitaire versus un SETAR stationnaire. Il est aussi démontré que ce test (supLR), dont la distribution asymptotique sous la nulle est obtenue analytiquement, est plus puissant que le Dickey-Fuller augmenté (ADF). Pour les séries stationnaires, il est possible de vérifier la linéarité de ces séries à l'aide du test de Hansen (1996) qui suppose que la série est stationnaire et  $\beta$ -mélangeant.

Le premier théorème parmi les trois proposés par Bec, Ben Salem et Carrasco (2004) offre les conditions nécessaires pour que  $y_t$ , un composant d'un processus Markov  $X_t$ , soit  $\beta$ -mélangeant avec une décroissance géométrique. Ainsi, sous les conditions du théorème,  $\{y_t\}$  est stationnaire  $\alpha$ - et  $\beta$ -mélangeant avec décroissance géométrique. Les deux autres théorèmes sont basés sur un modèle SETAR à trois régimes adapté pour l'analyse de taux de change réel :

$$\Delta y_t = \begin{cases} \alpha_{11}\Delta y_{t-1} + \dots + \alpha_{1p-1}\Delta y_{t-p+1} + \mu_1 + \rho_1 y_{t-1} + \sigma \varepsilon_t, & \text{if } y_{t-d} \leq -\lambda \\ \alpha_{21}\Delta y_{t-1} + \dots + \alpha_{2p-1}\Delta y_{t-p+1} + \mu_2 + \rho_2 y_{t-1} + \sigma \varepsilon_t, & \text{if } |y_{t-d}| \leq \lambda \\ \alpha_{31}\Delta y_{t-1} + \dots + \alpha_{3p-1}\Delta y_{t-p+1} + \mu_3 + \rho_3 y_{t-1} + \sigma \varepsilon_t, & \text{if } y_{t-d} \geq \lambda \end{cases}$$

Dans le modèle ci-dessus, la valeur de la variable  $d$  est fixée à 1 en suivant des études antérieures. Le deuxième et le troisième théorème font partie de l'étude des propriétés asymptotiques des tests. Les tests considérés sont wald, multiplicateur de lagrange et celui du rapport de vraisemblance. Dans le deuxième théorème les auteurs considèrent le cas quand le paramètre  $\lambda$  est choisi *a priori*. Ce théorème établit la distribution asymptotique de tests sous l'hypothèse nulle. Dans le cas du troisième théorème, le paramètre  $\lambda$  est inconnu. Ainsi, ce théorème établit les sup tests qui convergent en distribution vers une statistique avec une distribution pivotale sous l'hypothèse nulle. Une implication importante de ce dernier est la possibilité d'utiliser les valeurs critiques empiriques.

Afin de trouver les valeurs critiques empiriques des tests, Bec, Ben Salem et Carrasco (2004) proposent que la valeur de paramètre  $p$  soit 2 en s'appuyant sur leur étude empirique. Donc, le modèle ci-dessus prend la forme suivante :

$$\Delta y_t = \begin{cases} \alpha_1 \Delta y_{t-1} + \mu_1 + \rho_1 y_{t-1} + \sigma \varepsilon_t, & \text{if } y_{t-1} \leq -\lambda \\ \alpha_2 \Delta y_{t-1} + \mu_2 + \rho_2 y_{t-1} + \sigma \varepsilon_t, & \text{if } |y_{t-1}| \leq \lambda \\ \alpha_3 \Delta y_{t-1} + \mu_3 + \rho_3 y_{t-1} + \sigma \varepsilon_t, & \text{if } y_{t-1} \geq \lambda \end{cases}$$

C'est également le modèle sous l'hypothèse alternative ( $H_1$ ). Le modèle sous l'hypothèse nulle ( $H_0$ ) est  $\Delta y_t = \alpha \Delta y_{t-1} + \mu + \sigma \varepsilon_t$ . En se basant sur ces deux modèles, Bec, Ben Salem et Carrasco (2004) trouvent les valeurs critiques empiriques des trois tests sup, la taille et la puissance du test de racine unitaire (supLR). Les résultats montrent que la puissance du test supLR est plus grande que celle du test ADF. Ceci est reflété dans les résultats des tests appliqués aux données.

Les auteurs considèrent deux groupes de taux de change : cinq par rapport au dollar américain et six par rapport au mark allemand. Ils trouvent que parmi ces onze séries, cinq sont stationnaires. Les gagnants sont toutes des devises européennes par rapport au mark allemand : le franc français (FRF), la lire italienne (ITL), le franc belge (BEF), le florin hollandais (NLG), la peseta espagnole (ESP). Cependant, l'ADF rejette la nulle seulement pour FRF. De plus, le test de linéarité sup wald rejette la linéarité avec un haut niveau de confiance pour ITL, NLG et ESP. Pour FRF et BEF la linéarité est rejetée seulement au niveau 14% et 15 % respectivement. Il faut remarquer que le modèle sous  $H_1$  nous permettrait de tester des restrictions plus strictes imposées *a priori* sur le modèle SETAR présenté dans l'article de Bec, Ben Salem et Carrasco (2004). De plus, les auteurs concluent que les forces liées à la politique unificatrice de

l'Europe peuvent expliquer le fait que le test supLR rejette la nulle de la racine unitaire pour cinq parmi les six devises européennes.

Pour la deuxième partie de ce résumé de la littérature nous considérons les trois articles suivants : Meese et Rogoff (1983), Mark (1995) et Molodtsova et Papell (2009). Dans les paragraphes qui suivent nous examinons le choix de modèle dans chaque article. Ensuite nous résumons la méthode d'évaluation hors échantillon des modèles examinés. Nous abordons ces deux sujets en considérant les trois articles parallèlement.

Dans leur article Molodtsova et Papell (2009) font des références pertinentes et informatives aux études antérieures dans le domaine de la prévision hors-échantillon. Ils identifient les articles clés de la recherche dans ce domaine. De plus, Molodtsova et Papell (2009) résument, en ordre chronologique, les conclusions importantes concernant la qualité de prévisions de taux de change. Le point de départ est la conclusion de Meese et Rogoff (1983) « pas de prévision à court terme », la suivante est celle de Mark (1995) « prévision à long terme, mais pas à court terme » et finalement Chen et Mark (1996) concluent qu'il n'y a « pas de prévision pour aucun horizon ». Cependant, Molodtsova et Papell (2009) présentent de l'évidence en faveur de la prévision du taux de change à l'horizon d'un mois pour 12 pays d'OCDE par rapport au dollar américain.

En général, dans la littérature concernant la prévision de taux de change, le modèle sans pouvoir de prévision est considéré la marche aléatoire sans dérive. Ainsi la performance des modèles hors échantillon est évaluée en comparaison avec la marche aléatoire. Meese et Rogoff (1983) choisissent d'analyser trois modèles structurels populaires dans les années soixante-dix : le modèle monétaire de Frenkel-Bilson à prix flexible, le modèle monétaire Dornbusch-Frankel à prix rigide et le modèle Hooper-Morton incluant les effets de compte courant. Les trois modèles sont inclus dans la forme quasi-réduite suivante :

$$s = a_0 + a_1(m - \dot{m}) + a_2(y - \dot{y}) + a_3(r_s - \dot{r}_s) + a_4(\pi^e - \dot{\pi}^e) + a_5\overline{TB} + a_6\dot{\overline{TB}} + u,$$

où  $s$  est le logarithme du prix des devises étrangères en terme de dollar américaine,  $m - \dot{m}$  est le logarithme du ratio d'offre de monnaie de deux pays,  $y - \dot{y}$  le logarithme du ratio de revenu réel de deux pays,  $r_s - \dot{r}_s$  est la différence de taux de change à court terme,  $\pi^e - \dot{\pi}^e$  est la différence d'inflation attendue à long terme et  $\overline{TB}$  et  $\dot{\overline{TB}}$  sont les balances accumulées du commerce des É.-U. et respectivement du pays étranger en question. Chacun des trois modèles peut être obtenu en spécifiant des coefficients égaux à zéro. Donc, Frenkel-Bilson aura  $a_4 = a_5 = a_6 = 0$  et Dornbusch-Frankel aura  $a_5 = a_6 = 0$ . Un aspect problématique ici est le fait que la variable qui représente l'offre relative de monnaie et

celle qui représente le revenu relatif sont théoriquement considérées exogènes, pourtant *de facto* elles sont endogènes. Ceci est une raison importante pour laquelle Meese et Rogoff (1983) incluent dans l'analyse le modèle VAR sans contraintes (il ne pose pas de contraintes sur l'exogénéité des deux variables). Il faut aussi remarquer que les auteurs considèrent des séries temporelles univariées. Leur choix est un long AR sans tendance, sans ajustement saisonnier et sans les différences de variables. Long AR est un AR sans contrainte dont le plus long retard est une fonction de la taille de l'échantillon.

L'article Molodtsova et Papell (2009) est axé sur la règle de Taylor. Notamment, le modèle développé à partir de variables fondamentales de la règle de Taylor est contrasté avec la performance de trois autres modèles basés sur les fondements de taux d'intérêt, les fondements monétaires et les fondements de la parité de pouvoir d'achat (PPA). Ainsi le modèle basé sur les variables fondamentales de la règle de Taylor est donné par:

$$\Delta s_{t+1} = \omega - \omega_{u\pi}\pi_t + \omega_{f\pi}\tilde{\pi}_t - \omega_{uy}y_t + \omega_{fy}\tilde{y}_t + \omega_q\tilde{q}_t - \omega_{ui}i_{t-1} + \omega_{fi}\tilde{i}_{t-1} + \eta_t,$$

où  $s$  est le logarithme du prix de devises étrangères en terme de dollar américain. Le côté droit de l'équation ci-dessus change en fonction de différents cas proposés par Molodtsova et Papell (2009). Si le taux de change réel ( $q$ ) est spécifié par la banque centrale étrangère qui vise à maintenir la PPA, le modèle sera asymétrique. Pour construire un modèle avec lissage, il faut inclure les retards de taux d'intérêt ( $i$ ). Molodtsova et Papell (2009) modélisent l'ajustement de taux d'intérêt à l'aide d'un AR(1). De plus, le modèle sera considéré homogène si les deux banques centrales réagissent de la même façon à un changement dans le taux d'inflation ( $\pi$ ) et l'écart de production ( $y$ ), et si le coefficient du processus de lissage est le même. Sous ces conditions, mais pas nécessairement celle d'écart de production, et en s'assurant que les deux banques centrales ont la même cible d'inflation et le même taux d'intérêt réel d'équilibre, la constante ( $\omega$ ) pourrait être exclue. Ainsi, Molodtsova et Papell (2009) considèrent toutes les 16 combinaisons de quatre options binaires ci-dessus (ex. symétrique et non-symétrique). Les résultats démontrent que le modèle hétérogène avec constante, symétrique et avec lissage performe le mieux.

La liaison clé dans cet article et la relation entre le taux d'intérêt et le taux de change. Molodtsova et Papell (2009) la présentent dans une question: Comment une augmentation de la différence de taux d'intérêt va-t-elle affecter le taux de change prévu? Une discussion s'ensuit sur la parité des taux d'intérêt non couverte (UIRP), dont une conclusion importante est qu'UIRP ne tient pas à court terme.

Il faut aussi remarquer que Molodtsova et Papell (2009) estiment un nombre important de modèles. Ceci crée une situation où les résultats de la performance de modèles peuvent être influencés par leur nombre.

Afin d'exclure cette possibilité, Molodtsova et Papell (2009) effectuent le test pour la capacité supérieure de la prévision (SPA).

Mark (1995) a aussi étudié le modèle basé sur les fondements monétaires sous la forme :

$$e_{t+k} + e_t = \alpha_k + \beta_k z_t + v_{t+k,t} ,$$

où  $z_t$  est la différence entre la valeur fondamentale ( $f_t$ ) et le logarithme de taux de change ( $e_t$ ) :

$$f_t = (m_t - m_t^*) - \lambda(y_t - y_t^*)$$

Mark (1995) utilise deux méthodes pour évaluer la performance de son modèle par rapport à une marche aléatoire sans dérive. La première méthode est simplement le ratio entre la racine de la moyenne de l'erreur carrée (RMSE) de prédiction du modèle proposé et le RMSE de la marche aléatoire sans dérive. L'auteur dénote cette statistique OUR/RW. Cette approche est aussi utilisée par Meese et Rogoff (1983) ; ils proposent d'employer l'erreur absolue moyenne (MAE) de prédiction dans les cas quand les taux de change suivent un processus stable Parentian non-normal avec variance infinie ou la distribution du taux de change permet une quantité importante de valeurs extrêmes. Une idée très intéressante offerte par Meese et Rogoff (1983) dans ce contexte est de comparer le MAE avec l'erreur moyenne (ME) de prévision afin de voir si le modèle en question sous- ou sur-prédit de manière systématique. L'idée est que si la ME est très petite par rapport à la MAE nous pourrions conclure que les valeurs négatives dans ME « annulent » les valeurs positives. De cette façon, il n'y a pas de dominance ni d'un côté ni de l'autre.

La deuxième méthode utilisée par Mark (1995) est celle proposée par Diebold et Mariano (1995). En fait, c'est une méthode plus rigoureuse qui permet de décider si la différence entre l'erreur de prédiction de deux modèles est statistiquement significative. Ce test utilise la statistique t qui suit sous l'hypothèse nulle (la différence entre les erreurs de prédiction de deux modèles est zéro) une loi asymptotiquement standard normale. Grace à sa simplicité, la statistique de Diebold et Mariano (1995) (DM) est devenue très populaire comme mesure de comparaison de prévisions. Molodtsova et Papell (2009) emploient aussi le test DM. Cependant, puisque les modèles considérés par eux sont emboîtés, la statistique DM devient moins appropriée. La solution est la procédure d'inférence développée par Clark et West (2007)(CW). La statistique CW prend en considération que sous l'hypothèse nulle (une marche aléatoire sans dérive) la moyenne de l'erreur de prédiction carré (MSPE) d'échantillon pour le modèle alternatif devrait être plus grande que celle de la marche aléatoire.



Finalement, Meese et Rogoff (1983) concluent que l'évaluation hors-échantillon est un critère important à considérer quand on veut tester la performance de modèle au niveau de prévision. Ils supposent que pour un échantillon plus grand il serait possible de modéliser la dynamique de taux de change avec une série temporelle univariée. Les auteurs essaient de donner une explication plausible pour la performance faible de modèles structurels, mais ils n'arrivent pas à une conclusion définitive. Comme mentionné avant, Mark (1995) et Molodtsova et Papell (2009) arrivent aux résultats positifs pour des horizons différents. Dans les deux articles, les auteurs observent que les rigidités nominales doivent être permises dans les modèles étudiés. Mark (1995) conclut qu'à long terme (16 trimestres) le bruit qui est omniprésent à court terme s'annule tout en permettant de révéler les mouvements systématiques de taux de change.

### III. LA DESCRIPTION DES DONNÉES

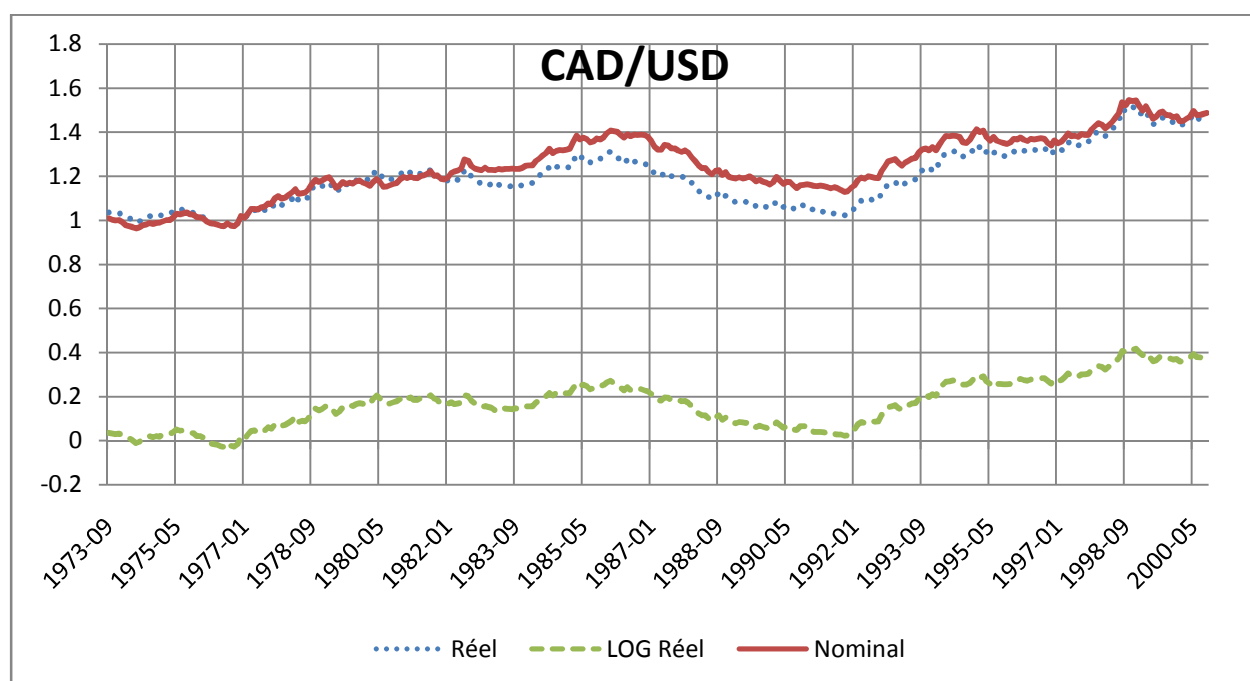
En suivant Bec, Ben Salem et Carrasco (2004) nous séparons les taux de change en deux groupes. Groupe A contient les taux de change de cinq devises par rapport au dollar américain(USD). Les cinq devises considérées sont : le franc français (FRF), le mark allemand (DEM), la lire italienne (ITL), la livre sterling (GBP) et le dollar canadien (CAD). Groupe B comprend les taux de change de six devises par rapport à DEM. Les six devises dans ce groupe sont : FRF, ITL, le franc belge (BEF), le florin hollandais (NLG), la peseta espagnole (ESP) et l'escudo portugais (PTE).

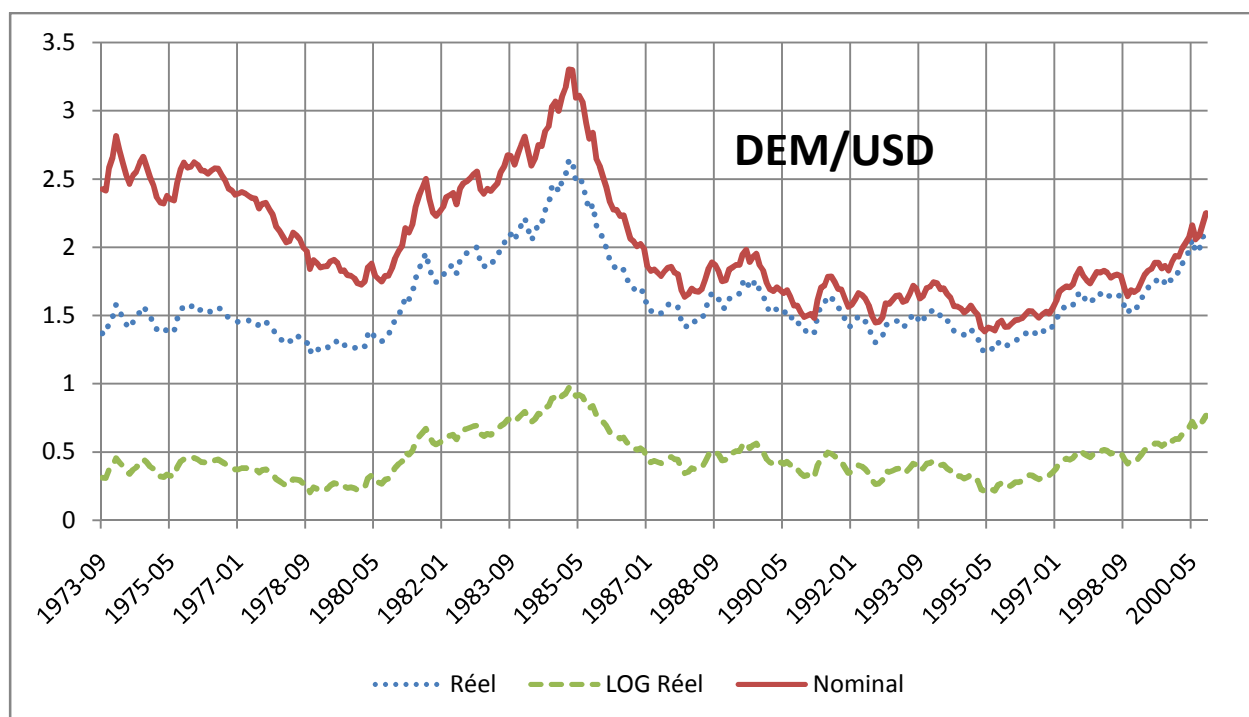
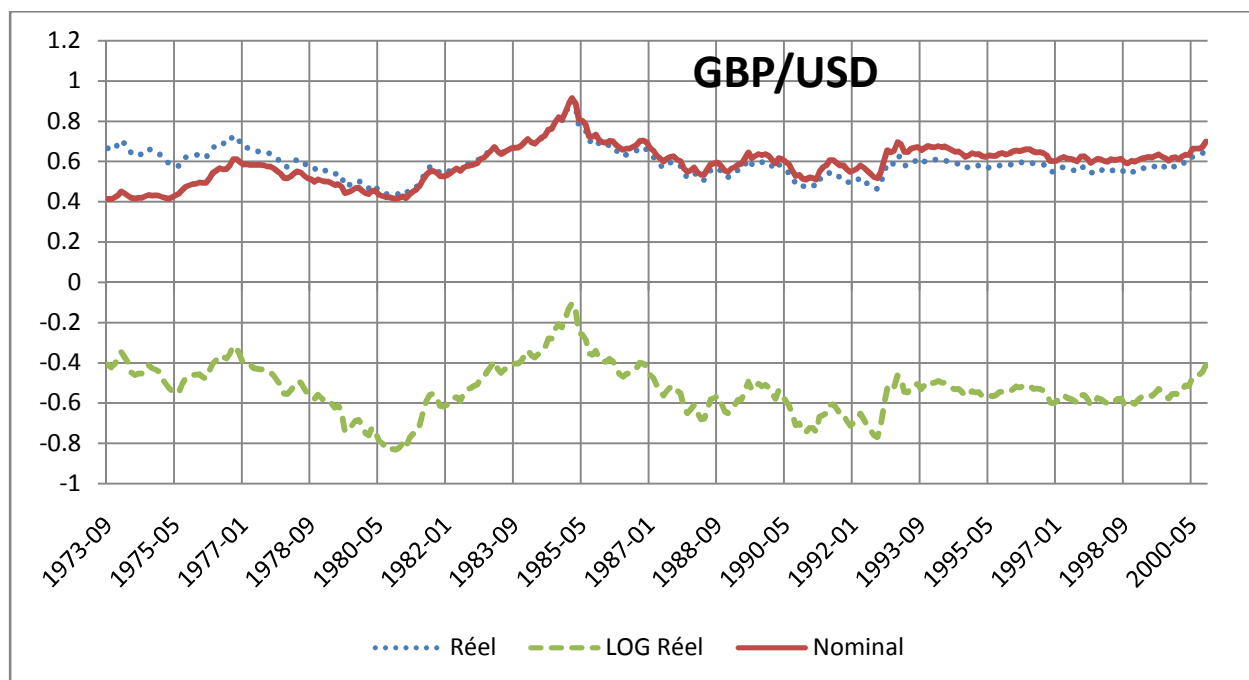
Afin de construire le taux de change réel à partir de la formule  $\epsilon_t = E_t * P_t^* / P_t$  ou sa version logarithmique  $\ln(\epsilon_t) = \ln(E_t) + \ln(P_t^*) - \ln(P_t)$ , nous trouvons les données pour le taux de change nominal ( $E_t$ ), soit le prix de la monnaie étrangère exprimé en monnaie domestique, et l'indice des prix à la consommation ( $P_t^*$  et  $P_t$ ) à partir de deux sources : la base de données économiques de la banque fédérale de St.Louis (FRED) et la base de données statistiques de l'Organisation de coopération et de développement économiques (OCDE). La FRED est la source des données sur le taux de change, alors que l'OCDE est la source des données sur l'indice des prix à la consommation (IPC). Il faut remarquer que les taux de change dans le groupe B ont été dérivés à partir de deux taux de change par rapport à USD. Autrement dit, l'USD était l'intermédiaire entre les deux devises. Ceci est une pratique souvent utilisée dans les travaux liés aux taux de change.

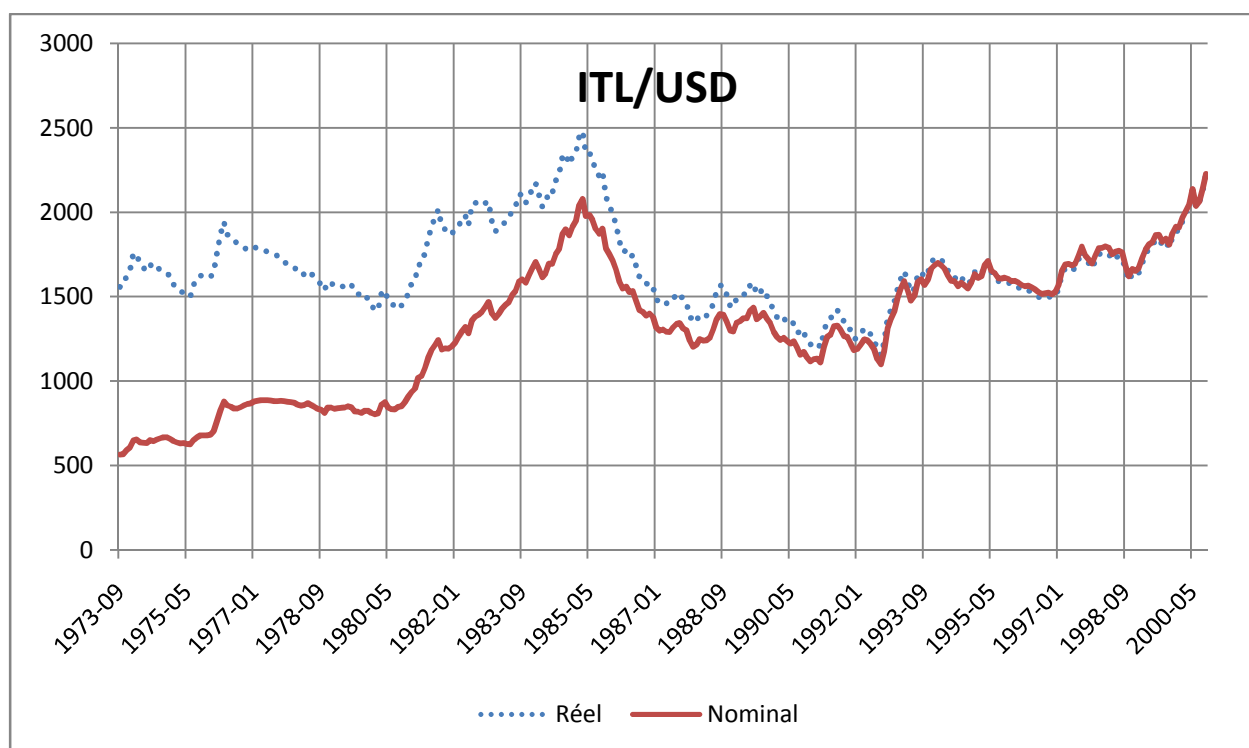
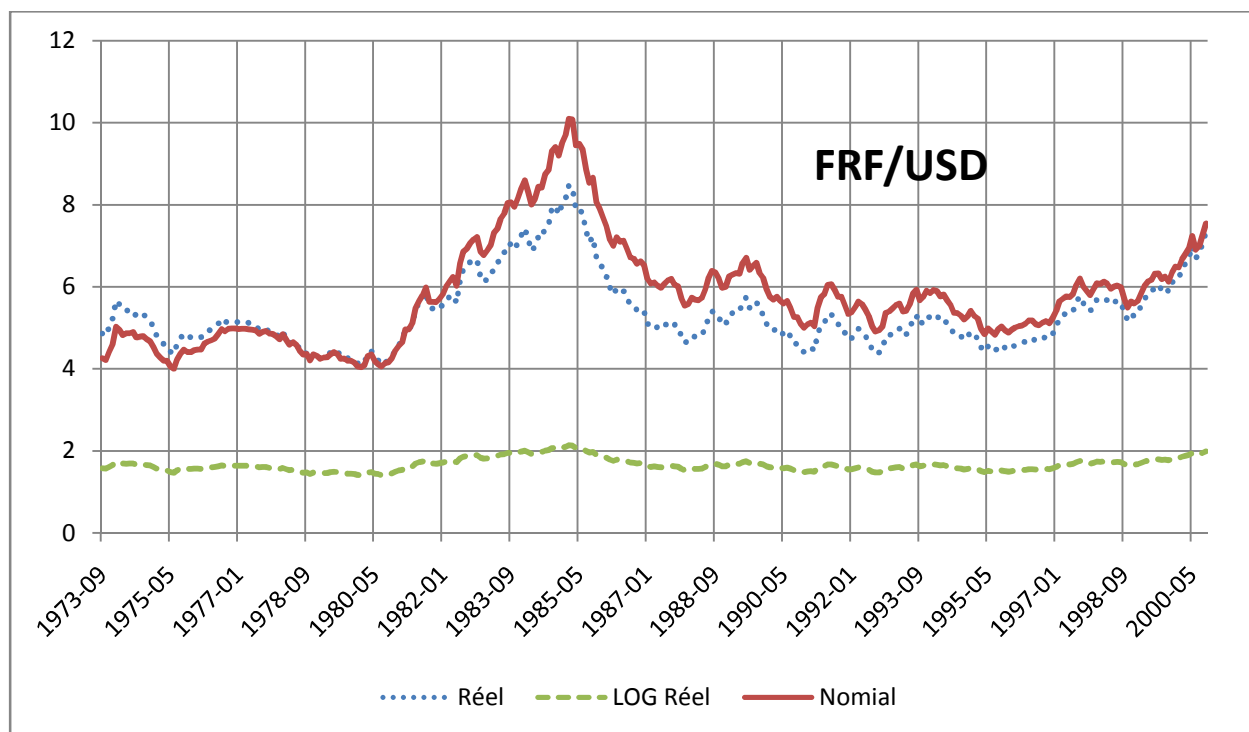
Les graphiques ci-dessous présentent la dynamique dans le temps des taux de change (nominal et réel) pour la période à partir de Septembre 1973 jusqu'à Septembre 2000. Les valeurs moyennes mensuelles

sont utilisées pour le taux de change nominal. Les premiers cinq graphiques représentent le groupe A et les six suivants le groupe B. Ils servent à nous donner une idée de l'évolution du taux de change. Il est aussi possible d'observer la relation entre le taux de change nominal et le taux de change réel. Le logarithme naturel du taux de change réel n'était pas inclus dans le graphique dans les cas où l'échelle ne permet pas d'observer les fluctuations dans le logarithme naturel.

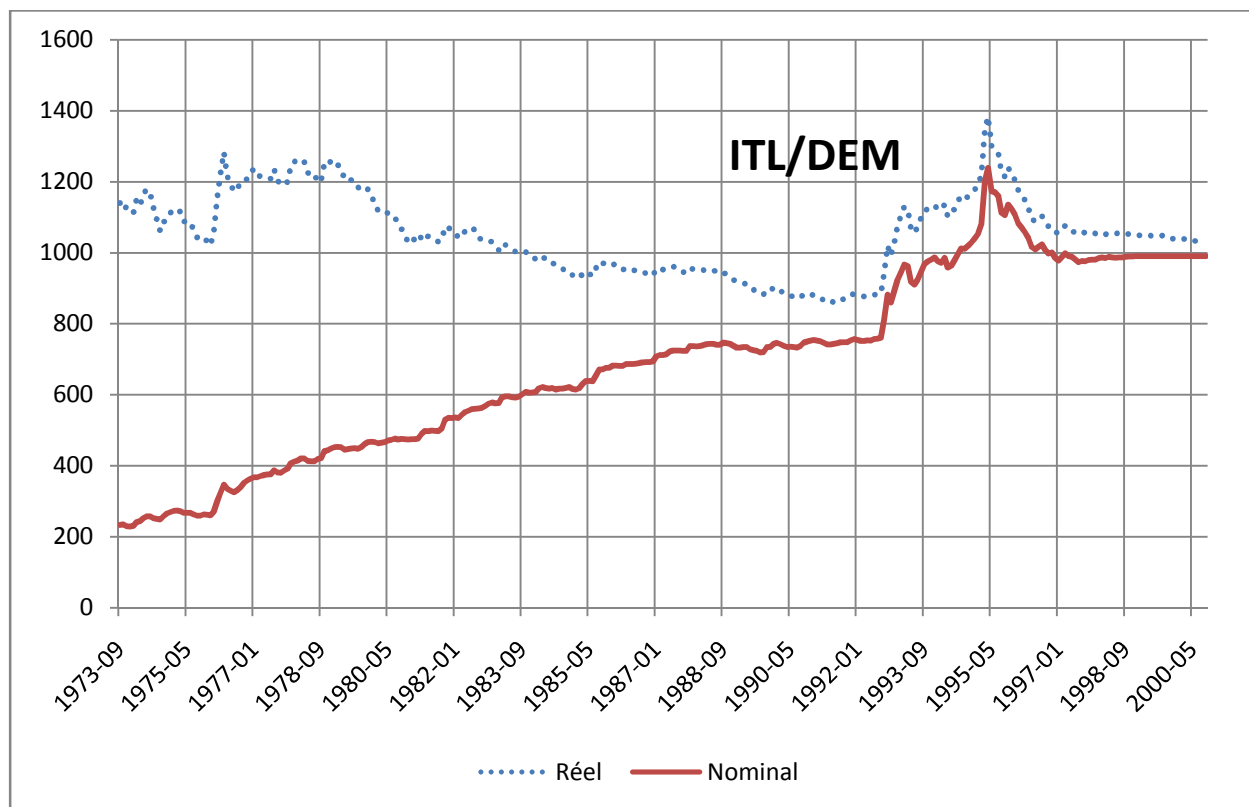
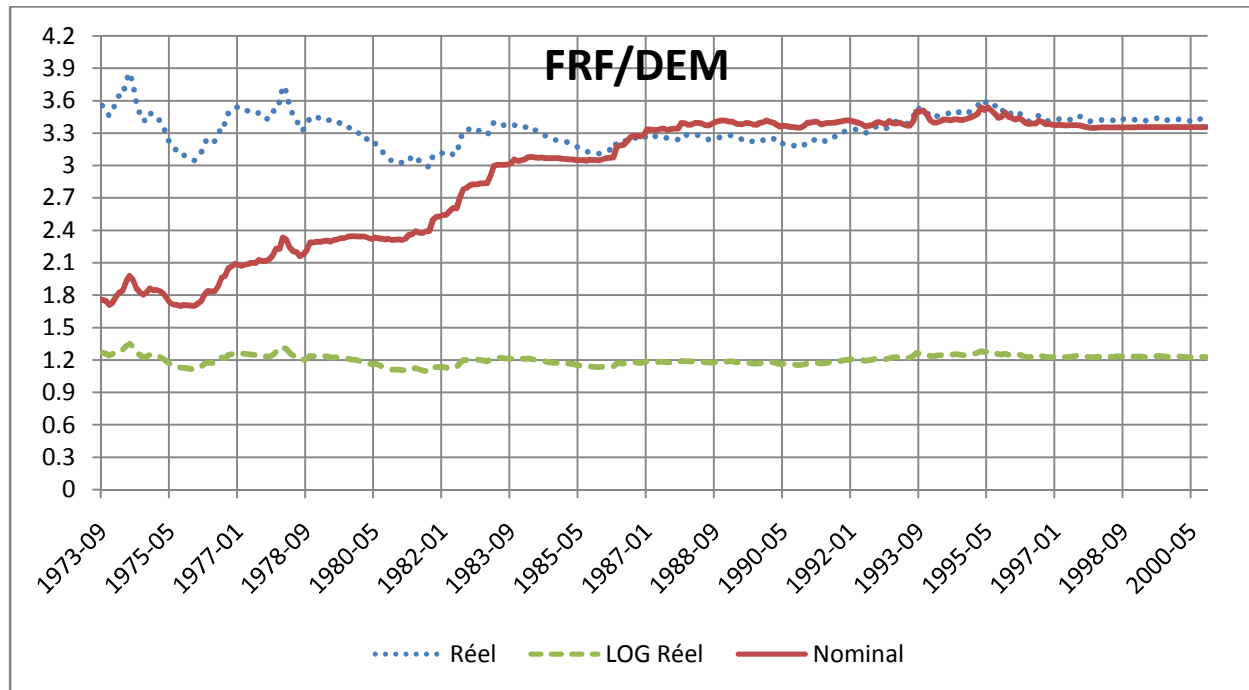
#### Groupe A

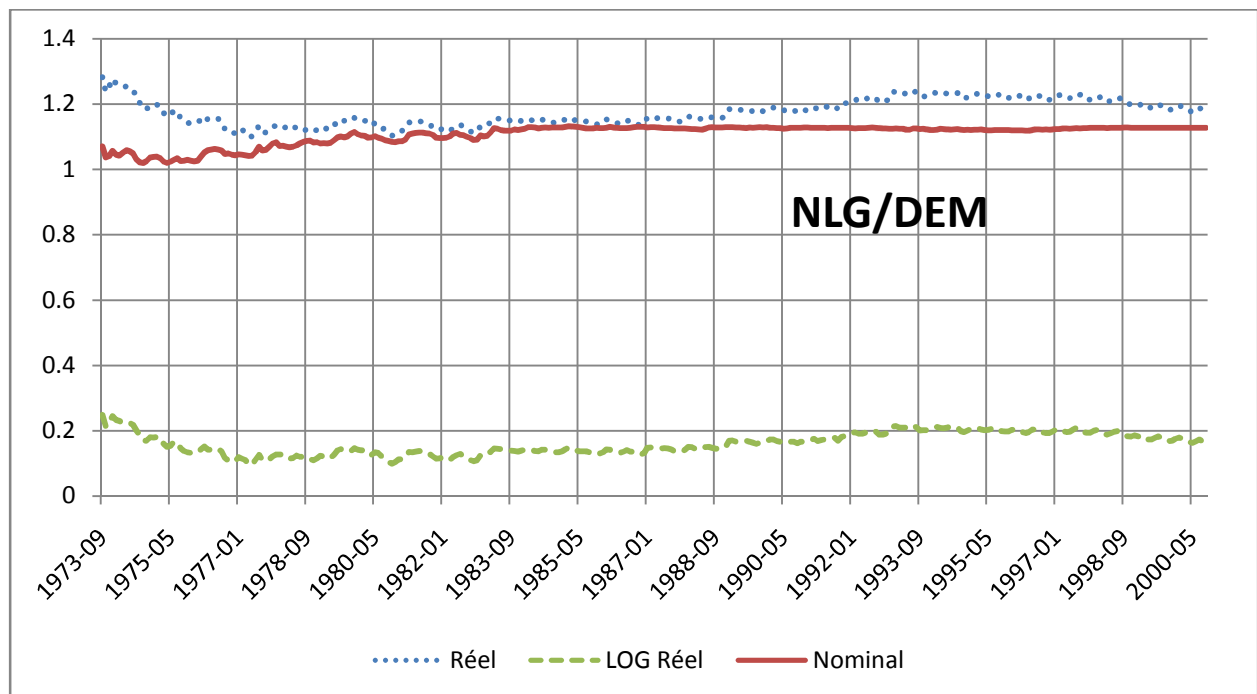
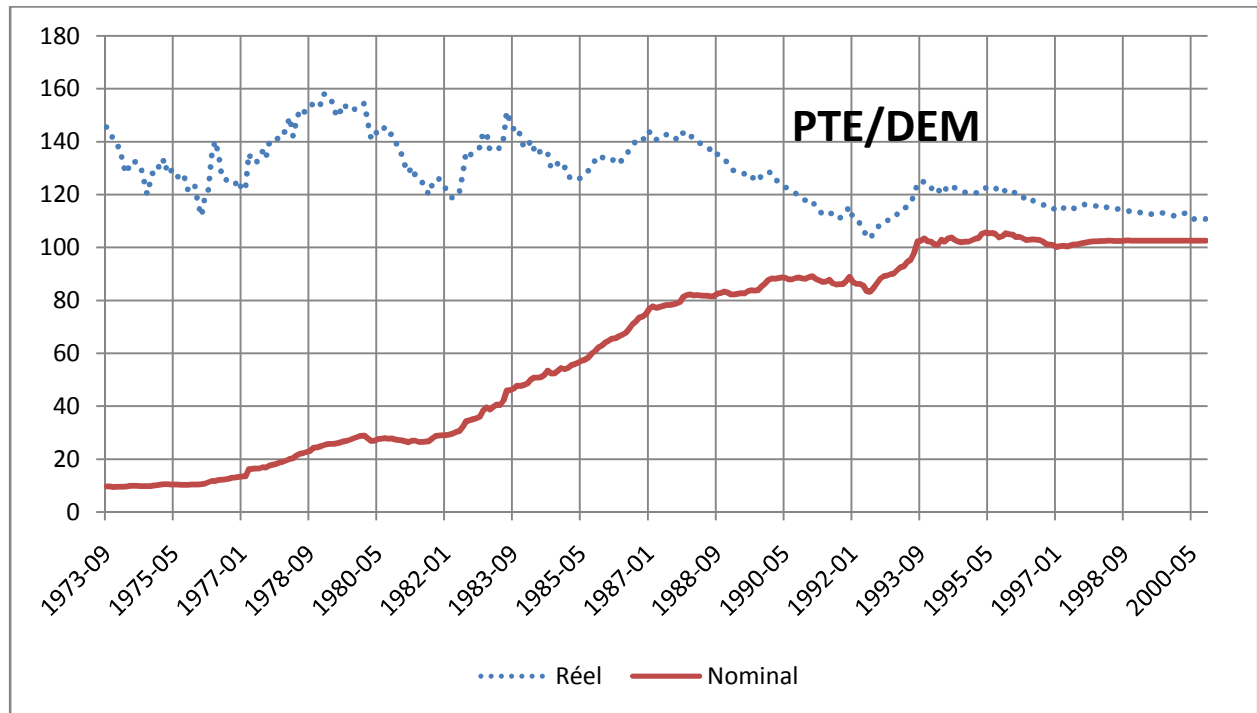


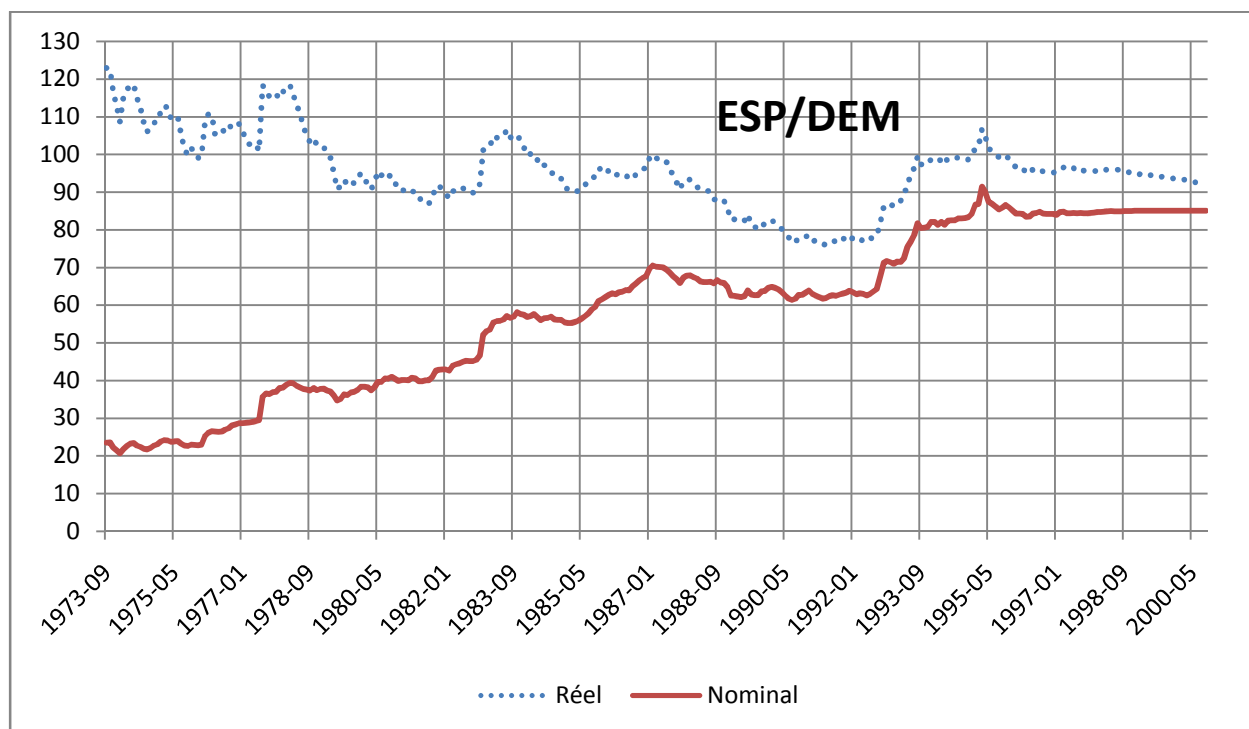
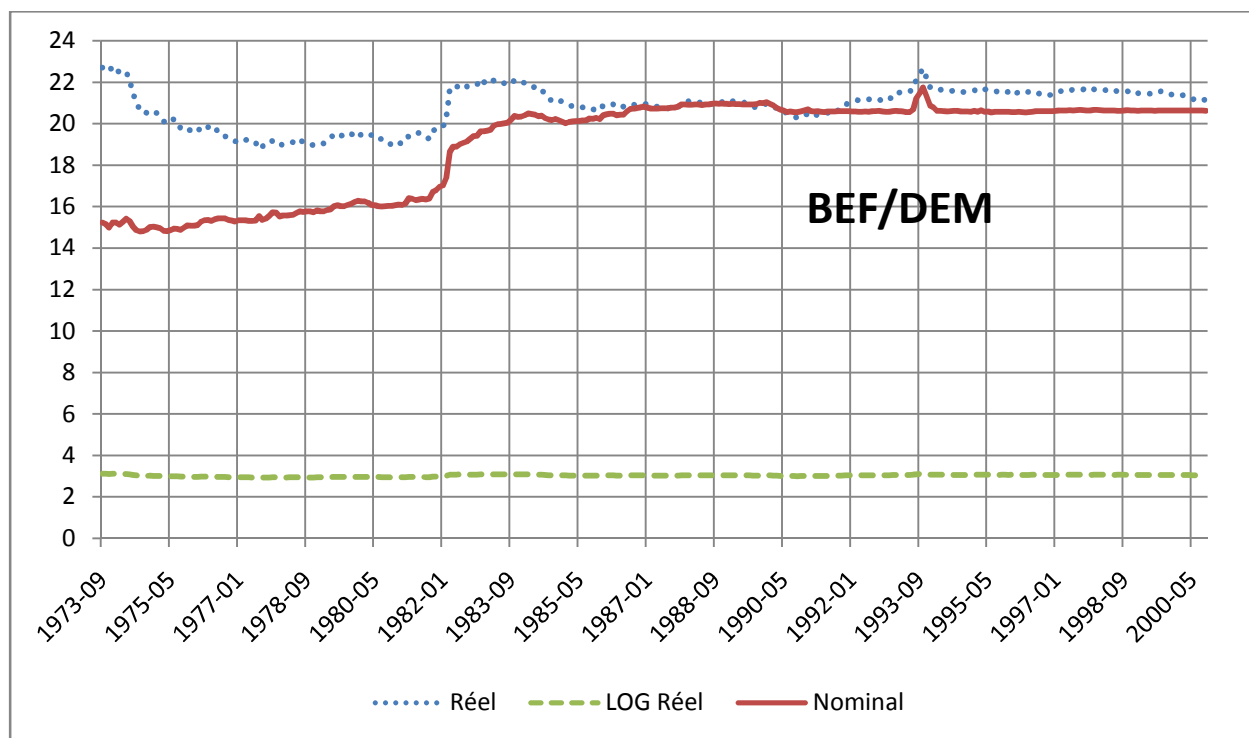




Groupe B







À partir de ces graphiques nous pourrions observer que les déviations majeures du taux de change réel se sont réduites au cours de la période. Il y a au moins deux facteurs qui peuvent illuminer cette situation. Premièrement, cette réduction serait due à l'intensification du commerce international grâce à

l'élimination des barrières commerciales. Deuxièmement, nous pouvons supposer que les banques centrales ont dû changer leurs politiques envers le taux de change et qu'elles sont devenues actives dans le marché des devises étrangères. Étant donné que cette tendance coïncide pour plusieurs pays, il est plus probable que la première raison est la principale.

Une question clé dans cette analyse est la stationnarité du taux de change réel. Un test de base est le test Dickey-Fuller augmenté (ADF). Les résultats de ce test sont présentés dans les deux tableaux ci-dessous.

Pour chaque devise, le retard est choisi à partir d'un de deux critères : AIC ou HQIC. Si les deux valeurs de retards obtenus de deux critères diffèrent, nous choisissons celle pour laquelle la valeur absolue de la statistique de test ADF est plus élevée. La p-valeur incluse dans les tableaux est la p-valeur approximée de MacKinnon.

Tableau 1

(/USD)	CAD	GBP	DEM	FRF	ITL
Nombre de retards (différence)	1	3	1	3	1
Valeur de la statistique	-0.464	-2.678	-1.652	-1.604	-1.795
p-valeur	0.8989	0.0779	0.4561	0.4816	0.3829

Tableau 2

(/DEM)	FRF	ITL	PTE	NLG	BEF	ESP
Nombre de retards (différence)	1	3	1	1	3	1
Valeur de la statistique	-3.102	-1.829	-1.554	-1.928	-2.251	-1.788
p-valeur	0.0264	0.3664	0.5066	0.3188	0.1883	0.3865

Nous observons que dans le groupe A nous ne pouvons pas rejeter l'hypothèse nulle (la présence de racine unitaire) pour toutes les devises. Une exception marginale est GBP. Dans ce cas nous pouvons rejeter l'hypothèse nulle seulement au niveau de confiance de 10%. Quant au groupe B, il y a une seule devise, soit FRA, pour laquelle nous pouvons rejeter l'hypothèse nulle avec confiance (5% et 10%).



## IV. L'ESTIMATION ET LA PRÉVISION DU MODÈLE SETAR À TROIS RÉGIMES

La forme du modèle SETAR utilisé dans cette analyse est celle donnée par Bec, Ben Salem et Carrasco (2004).

$$\Delta y_t = \begin{cases} \alpha_{11}\Delta y_{t-1} + \dots + \alpha_{1p-1}\Delta y_{t-p+1} + \mu_1 + \rho_1 y_{t-1} + \sigma \varepsilon_t, & y_{t-d} \leq -\lambda \\ \alpha_{21}\Delta y_{t-1} + \dots + \alpha_{2p-1}\Delta y_{t-p+1} + \mu_2 + \rho_2 y_{t-1} + \sigma \varepsilon_t, & |y_{t-d}| < \lambda \\ \alpha_{31}\Delta y_{t-1} + \dots + \alpha_{3p-1}\Delta y_{t-p+1} + \mu_3 + \rho_3 y_{t-1} + \sigma \varepsilon_t, & y_{t-d} \geq \lambda \end{cases}$$

Les auteurs suggèrent  $p = 2$  en s'appuyant sur des études empiriques. De plus, la valeur du paramètre  $d$  est fixée à 1. Ce choix est motivé surtout par la rigidité des prix. Ainsi nous obtenons le modèle ci-dessous, qui est aussi celui que nous utilisons pour les régressions. Pour l'instant, nous n'imposons aucune restriction sur les coefficients.

$$\Delta y_t = \begin{cases} \alpha_1 \Delta y_{t-1} + \mu_1 + \rho_1 y_{t-1} + \sigma \varepsilon_t, & y_{t-1} \leq -\lambda \\ \alpha_2 \Delta y_{t-1} + \mu_2 + \rho_2 y_{t-1} + \sigma \varepsilon_t, & |y_{t-1}| < \lambda \\ \alpha_3 \Delta y_{t-1} + \mu_3 + \rho_3 y_{t-1} + \sigma \varepsilon_t, & y_{t-1} \geq \lambda \end{cases} \quad (1)$$

où  $\varepsilon_t$  est iid  $N(0,1)$ . À l'aide de variables indicatrices, le modèle ci-dessus est transformé dans la forme (2) qui peut être estimée par la méthode des moindres carrés ordinaires (MCO).

$$\Delta y_t = \alpha_1 I_1 \Delta y'_{t-1} + \rho_1 I_1 y'_{t-1} + \alpha_2 I_2 \Delta y'_{t-1} + \rho_2 I_2 y'_{t-1} + \alpha_3 I_3 \Delta y'_{t-1} + \rho_3 I_3 y'_{t-1} + I_1 + I_2 + I_3 + \sigma \varepsilon_t, \quad (2)$$

où  $I_1, I_2$  et  $I_3$  sont les variables indicatrices correspondant aux trois régimes, respectivement:

$$y_{t-1} \leq -\lambda, |y_{t-1}| < \lambda \text{ et } y_{t-1} \geq \lambda.$$

### 1. L'estimation

Le logiciel statistique STATA, que nous utilisons, n'a pas une routine pour estimer les paramètres d'un modèle SETAR à trois régimes. Pour cette raison nous développons notre propre routine basée sur l'estimation par MCO du modèle (2), où le paramètre  $\lambda$  (seuil) est estimé par la méthode « inf » décrite ci-dessous. Afin de nous assurer que notre approche fonctionne bien, nous vérifions si les valeurs estimées par MCO et le seuil estimé convergent en probabilité vers leurs vraies valeurs.

Les simulations que nous faisons ont deux objectifs principaux. Premièrement, elles servent à vérifier le bon fonctionnement de notre routine d'estimation de modèle SETAR à trois régimes et deuxièmement elles nous donnent une meilleure idée de la taille d'échantillon nécessaire pour obtenir une valeur espérée du paramètre estimé autour de sa vraie valeur.

La méthode d'estimation « inf » est inspirée de Davies (1987) et Bec, Ben Salem et Carrasco (2004). Supposons un ensemble  $\Lambda$  qui contient toutes les valeurs discrètes possibles de seuil. Dans notre cas, par défaut, cet ensemble contient toutes les observations de  $y$  en valeur absolue en excluant les 10% des observations extrêmes de deux côtés. Les valeurs extrêmes sont exclues afin de réduire la probabilité que dans un parmi les trois régimes il n'y a presque aucune observation. Dans nos simulations nous suivons le nombre d'observations dans chaque régime afin de nous assurer que ce niveau de coupure (10%) est bien choisi.

Ainsi, l'estimateur de  $\lambda$  est donné par l'infimum des valeurs de la somme des carrés des résidus(SCR) obtenues en estimant le modèle (2) pour chaque valeur possible de seuil dans l'ensemble  $\Lambda$ .

$$\hat{\lambda} = \inf \{SCR(\lambda) : \lambda \in \Lambda\}$$

Donc, ici la valeur estimée de seuil ( $\hat{\lambda}$ ) est celle qui correspond à la SCR la plus petite.

Les propriétés statistiques de l'estimateur MCO développées analytiquement servent comme point de référence pour l'évaluation de notre approche d'estimation d'un modèle SETAR. Étant donné que dans le modèle considéré les variables explicatives sont des variables retardées, nous ne pouvons pas imposer l'exogénéité stricte, une condition indispensable pour l'absence de biais de MCO. Cependant, il est possible d'imposer l'exogénéité contemporaine, soit  $E(\varepsilon_t | x_t) = 0$ , où  $x_t = (y_{t-1}, y_{t-2})$ . Sous l'exogénéité contemporaine, l'estimateur MCO converge en probabilité vers la vraie valeur de paramètre. C'est-à-dire qu'avec l'augmentation de l'échantillon la valeur espérée des paramètres estimés s'approche de la vraie valeur et la variance diminue.

Dans le contexte de simulations, la valeur espérée pour les paramètres estimés est donnée par la moyenne échantillonnale. À cet effet, nous procédons de la façon suivante: pour chaque répétition nous obtenons une valeur estimée de paramètre, ensuite toutes les valeurs estimées obtenues sont sauvegardées et donc nous obtenons un échantillon de taille égale au nombre de répétitions. Ainsi, la moyenne échantillonnale est calculée. Ensuite, la variance et l'écart-type sont calculés.

Les tableaux 3,4 et 5 contiennent les résultats des simulations. Chaque simulation ou expérience est dénotée par la taille d'échantillon(T) et par le nombre de répétitions(R) Monte-Carlo et chacune est décrite par trois colonnes. Par exemple, « T=1000, R=1000 » décrit une expérience pour une taille d'échantillon de 1000 observations et un nombre de répétitions égal à 1000. La première colonne(1) contient la valeur espérée de paramètres estimés donnée par la moyenne échantillonnale, la deuxième(2) l'écart-type échantillonnale et la troisième le ratio t sous l'hypothèse nulle que la valeur de paramètre estimé est 0. Par défaut, nous considérons le niveau de confiance de 5% pour lequel la valeur critique

de  $t_c \approx 1,96$ . Donc, si la valeur absolue des valeurs dans la troisième colonne d'une de simulations est plus grande que 1.96 nous pouvons rejeter l'hypothèse nulle.

Les six dernières lignes dans chacun des tableaux sont lues différemment de celles qui les précèdent; il s'agit des proportions de données dans chacun de trois régimes. Les termes sim1, sim2 et sim3 dénotent la proportion moyenne d'observations dans chaque régime lors de la simulation du PGD, tandis que est1, est2 et est3 dénotent la proportion moyenne d'observations lors de l'estimation. Donc, la comparaison est faite entre sim1, sim2, sim3 et est1, est2, est3.

En considérant un processus générateur de données (PGD) fondé sur le modèle SETAR à trois régimes, nous nous attendons à avoir dans les petits échantillons des valeurs espérées de paramètres estimés qui sont différentes de leurs vraies valeurs. Notamment, il s'agit de biais où  $E(\hat{\theta}) - \theta \neq 0$ . Toutefois, quand nous augmentons la taille de l'échantillon nous nous attendons à observer des valeurs espérées de paramètres estimés qui s'approchent des vraies valeurs. En fait, le biais converge vers zéro avec l'augmentation de la taille d'échantillon.

Nous examinons aussi comment la taille d'échantillon et la valeur de paramètre  $\lambda$  influencent la vitesse de convergence en probabilité de paramètres estimés vers les vraies valeurs. Nous avons trouvé que 1000 répétitions suffisent pour établir si les paramètres estimés s'approchent ou non de la vraie valeur de paramètres pour une taille donnée d'un échantillon. C'est-à-dire que si la valeur espérée du paramètre estimé converge en probabilité vers sa vraie valeur, la valeur espérée obtenue à partir de 1000 répétition est autour de la vraie valeur. En augmentant le nombre de répétitions, l'information que nous obtenons est marginale, c.-à-d. que les valeurs espérées de paramètres estimés varient très peu lors d'une augmentation de nombre de répétitions au dessus de 1000. Par exemple considérons le tableau 4 et la taille d'échantillon de 2000, la valeur espérée de  $\hat{\alpha}_3$  est déjà autour de la vraie valeur  $\alpha_3$  quand le nombre de répétitions est 1000. Quand le nombre de répétition augmente de 1000 à 3000 ou de 1000 à 5000, la différence entre les valeurs espérées de  $\hat{\alpha}_3$  est négligeable. En considérant ce résultat, nous nous limitons à 1000 répétitions pour les simulations présentés dans le tableau 5.

La performance de la méthode d'estimation est une fonction croissante du nombre de données disponibles dans le régime en question. Ainsi, la plus haute vitesse de convergence correspond au régime contenant le plus d'observations. Bien sûr que plus la distribution d'observations parmi les régimes est homogène, moins est remarquable la différence entre leurs vitesses de convergence. Par implication, la taille de d'échantillon est cruciale pour la convergence et évidemment pour la vitesse de convergence.

Quant à l'impact de la taille de seuil( $\lambda$ ) sur la vitesse de convergence, nous observons à partir de tableaux 3 et 4 qu'une augmentation de 0,026 à 0,04 n'a pas un impact important sur la vitesse de convergence (c.-

à-d. pour un  $R$  donné, nous regardons quels paramètres constituant un ensemble parmi les ensembles correspondant aux deux valeurs de  $\lambda$  sont plus proches des vraies valeurs quand le  $T$  est augmenté). Dans tableau 5,  $\lambda$  est trop petit par rapport aux valeurs générées à partir de paramètres donnés ( $\lambda = 0,01$ ); il est notamment instructif de le comparer à la valeur de  $\sigma = 0,07$ . Ainsi, la proportion d'observations dans le régime de milieu est basse. Même quand nous laissons tomber la coupure de 10% de valeurs extrêmes (expériences [4] et [5] dans le tableau 5), ce qui permet au  $\hat{\lambda}$  de prendre des valeurs extrêmes, la proportion des observations dans le régime de milieu (2) lors de l'estimation est loin de la proportion lors du PGD. Par conséquent,  $\hat{\lambda}$  est loin de sa vraie valeur même pour la taille d'échantillon de 5000 observations – la plupart des observations sont dans le premier régime ce qui entraîne une forte vitesse de convergence dans le premier régime et une faible vitesse de convergence dans les deux autres.

L'aspect le plus important dans l'évaluation de la routine d'estimation est la convergence des paramètres estimés vers les vraies valeurs et la diminution de l'écart-type dès que la taille d'échantillon augmente. Ceci est aussi notre critère de base pour l'évaluation de la routine d'estimation du modèle (2). Dans les deux dernières expériences ([6] et [7]) présentées dans le tableau 3, nous considérons respectivement le cas d'une racine unitaire dans le régime de milieu et le cas d'une racine unitaire dans le régime 1. Dans le modèle (1), si nous imposons une racine unitaire dans le régime de milieu,  $\rho_2$  sera 0 et si nous imposons une racine unitaire dans le régime 1,  $\rho_1$  sera 0.

Dans l'expérience [5] de tableau 3, les valeurs espérées des paramètres estimés sont pratiquement égales aux vraies valeurs. De plus, nous observons la convergence en distribution en comparant les écarts-type des expériences [3] et [5]. Comme dans l'expérience [5] la taille d'échantillon est plus grande que celle de l'expérience [3], les écart-types baissent. En imposant une racine unitaire locale dans le régime de milieu (l'expérience [6]), nous remarquons que les résultats sont très similaires aux ceux de l'expérience [5]. Donc, une racine unitaire(RU) dans le régime de milieu n'a pas un impact important ni sur la valeur de paramètres estimés, ni sur la convergence en probabilité. Pourtant, en présence d'une RU dans le régime de milieu, beaucoup moins d'observations se trouvent dans ce régime, soit 15% par rapport au 50% dans [5].

Pourtant, quand la racine unitaire est imposée dans le régime 1 (l'expérience [7]), l'estimateur MCO ne converge pas vers les vraies valeurs dans les deux autres régimes. La proportion moyenne d'observations dans le régime 1 est 92% et les deux autres régimes partagent seulement 8% d'observations. De son côté, le paramètre de seuil n'est pas identifié. Donc, une RU dans le régime 1 affecte défavorablement les résultats de l'estimation, tandis qu'une RU dans le régime de milieu n'affecte pas d'une façon importante les bons résultats obtenus dans l'expérience [5]. Nous prenons ceci en considération pour les expériences présentées dans les tableaux 4 et 5.

Tableau 3

Vraies Valeurs		[1]T=1000, R=1000			[2]T=1000, R=3000			[3]T=2000, R=1000			[4]T=2000, R=3000		
		E()	Écart-type	Ratio t	E()	Écart-type	Ratio t	E()	Écart-type	Ratio t	E()	Écart-type	Ratio t
$\sigma$	<b>0,007</b>	0,0070	0,0002	44,3158	0,0070	0,0002	44,3016	0,0070	0,0001	61,9097	0,0070	0,0001	63.2541
$\alpha 1$	<b>-0,043</b>	2,9359	86,4307	0,0340	-0,0093	0,8075	-0,0115	-0,0207	0,3536	-0,0585	-0,0318	0,1948	-0.1630
$\alpha 2$	<b>0,202</b>	0,1989	0,0660	3,0137	0,2003	0,0607	3,3014	0,2026	0,0328	6,1847	0,2030	0,0358	5.6698
$\alpha 3$	<b>0,220</b>	0,2299	0,1365	1,6844	0,2383	0,2323	1,0257	0,2210	0,0720	3,0699	0,2294	0,1426	1.6088
$\rho 1$	<b>-0,075</b>	2,1129	93,5131	0,0226	-0,1736	2,7150	-0,0639	-0,0377	1,6171	-0,0233	-0,0977	0,3855	-0.2533
$\rho 2$	<b>-0,240</b>	-0,2379	0,0574	-4,1450	-0,2388	0,0544	-4,3898	-0,2410	0,0301	-8,0161	-0,2407	0,0318	-7.5582
$\rho 3$	<b>-0,126</b>	-0,1685	0,2483	-0,6786	-0,1711	0,6482	-0,2640	-0,1396	0,0529	-2,6407	-0,1397	0,2801	-0.4987
$\mu 1$	<b>-0,005</b>	0,1098	4,2398	0,0259	-0,0070	0,0722	-0,0970	-0,0038	0,0463	-0,0819	-0,0054	0,0108	-0.5032
$\mu 2$	<b>0,001</b>	0,0008	0,0013	0,6240	0,0008	0,0014	0,5880	0,0010	0,0005	1,9232	0,0010	0,0004	2.7760
$\mu 3$	<b>0,005</b>	0,0065	0,0087	0,7434	0,0064	0,0187	0,3414	0,0054	0,0021	2,5490	0,0054	0,0112	0.4827
$\lambda$	<b>0,026</b>	0,0274	0,0073	3,7512	0,0270	0,0063	4,3072	0,0262	0,0017	15,5289	0,0262	0,0020	12.8160
sim1		0,2993	0,2358		0,2892	0,2392		0,3144	0,1955		0,3178	0,1954	
sim2		0,5430	0,1861		0,5520	0,1881		0,5333	0,1523		0,5295	0,1511	
sim3		0,1561	0,0786		0,1571	0,0830		0,1513	0,0650		0,1518	0,0655	
est1		0,3189	0,2455		0,3139	0,2579		0,3338	0,2092		0,3365	0,2085	
est2		0,5300	0,1927		0,5317	0,2007		0,5173	0,1626		0,5141	0,1610	
est3		0,1511	0,0868		0,1544	0,0929		0,1489	0,0683		0,1494	0,0692	

Vraies Valeurs		[5]T=5000, R=1000			[6]T=5000, R=1000 (RU régime 2)			[7]T=5000, R=1000 (RU régime 1)		
		E()	Écart-type	Ratio t	E()	Écart-type	Ratio t	E()	Écart-type	Ratio t
$\sigma$	<b>0,007</b>	0,0070	0,0001	102,4641	0,0070	0,0001	102,1766	0,0070	0,0001	99.4284
$\alpha 1$	<b>-0,043</b>	-0,0414	0,0258	-1,6044	-0,0428	0,0186	-2,2953	-0,0437	0,0265	-1.6503
$\alpha 2$	<b>0,202</b>	0,2024	0,0196	10,3083	0,1980	0,0402	4,9233	0,0362	0,1007	0.3593
$\alpha 3$	<b>0,220</b>	0,2219	0,0382	5,8020	0,2217	0,0354	6,2564	0,0494	0,1854	0.2666
$\rho 1$	<b>-0,075</b>	-0,0757	0,0109	-6,9430	-0,0761	0,0077	-9,8939	<b>0,0000</b>	<b>0,0002</b>	<b>-0.0600</b>
$\rho 2$	<b>-0,240</b>	-0,2405	0,0138	-17,4318	<b>0,0007</b>	<b>0,0284</b>	<b>0,0243</b>	-0,0367	0,0864	-0.4249
$\rho 3$	<b>-0,126</b>	-0,1305	0,0218	-5,9993	-0,1283	0,0302	-4,2565	-0,0355	0,1885	-0.1883
$\mu 1$	<b>-0,005</b>	-0,0050	0,0008	-6,4223	-0,0051	0,0005	-9,4082	-0,0053	0,0049	-1.0874
$\mu 2$	<b>0,001</b>	0,0010	0,0002	6,5481	0,0009	0,0005	1,8660	-0,0026	0,0025	-1.0643
$\mu 3$	<b>0,005</b>	0,0052	0,0010	5,1899	0,0051	0,0015	3,3635	0,0016	0,0133	0.1175
$\lambda$	<b>0,026</b>	0,0260	0,0002	119,3489	0,0273	0,0034	8,0752	6,4083	7,2816	0.8801
sim1		0,3430	0,1346		0,5845	0,1398		0,9181	0,0867	
sim2		0,5093	0,1053		0,1516	0,0460		0,0628	0,0654	
sim3		0,1473	0,0418		0,2635	0,0988		0,0187	0,0239	
est1		0,3515	0,1386		0,5934	0,1390		0,6889	0,2646	
est2		0,5025	0,1084		0,1542	0,0448		0,3030	0,2670	
est3		0,1459	0,0428		0,2524	0,1032		0,0081	0,0219	

Tableau 4 ( $\rho_2 = 0$  et  $\lambda = 0,04$ )

Vraies Valeurs		[1]T=1000, R=1000			[2]T=1000, R=3000			[3]T=2000, R=1000		
		E()	Écart-type	Ratio t	E()	Écart-type	Ratio t	E()	Écart-type	Ratio t
$\sigma$	0,007	0,0070	0,0002	44,3706	0,0070	0,0002	44,1061	0,0070	0,0001	62.1815
$\alpha_1$	-0,043	0,0062	0,8950	0,0070	0,0855	2,3588	0,0363	-0,0197	0,1730	-0.1138
$\alpha_2$	0,202	0,2075	0,0623	3,3299	0,2091	0,0629	3,3223	0,2046	0,0376	5.4447
$\alpha_3$	0,220	0,2239	0,0691	3,2402	0,2256	0,0708	3,1861	0,2226	0,0438	5.0793
$\rho_1$	-0,075	-0,1636	1,1270	-0,1452	-0,2321	2,2126	-0,1049	-0,1071	0,2246	-0.4767
$\rho_2$	0,000	-0,0150	0,0473	-0,3173	-0,0163	0,0516	-0,3169	-0,0047	0,0204	-0.2294
$\rho_3$	-0,126	-0,1248	0,0634	-1,9694	-0,1275	0,0821	-1,5529	-0,1244	0,0387	-3.2127
$\mu_1$	-0,005	-0,0087	0,0529	-0,1643	-0,0098	0,0717	-0,1364	-0,0062	0,0094	-0.6627
$\mu_2$	0,001	0,0011	0,0007	1,5763	0,0011	0,0006	1,7557	0,0011	0,0003	3.1766
$\mu_3$	0,005	0,0047	0,0037	1,2869	0,0049	0,0050	0,9811	0,0048	0,0021	2.2776
$\lambda$	0,040	0,0366	0,0098	3,7415	0,0361	0,0101	3,5914	0,0387	0,0058	6.7077
siml1		0,1626	0,1525		0,1547	0,1509		0,1642	0,1160	
siml2		0,5122	0,0912		0,5148	0,0898		0,5080	0,0676	
siml3		0,3235	0,0919		0,3288	0,0930		0,3270	0,0713	
estl1		0,1754	0,1647		0,1680	0,1645		0,1701	0,1225	
estl2		0,4494	0,1712		0,4432	0,1746		0,4833	0,1154	
estl3		0,3752	0,1914		0,3889	0,1995		0,3466	0,1242	
Vraies Valeurs		[4]T=2000, R=3000			[5]T=2000, R=5000			[6]T=5000, R=1000		
		E()	Écart-type	Ratio t	E()	Écart-type	Ratio t	E()	Écart-type	Ratio t
$\sigma$	0,007	0,0070	0,0001	61,4881	0,0070	0,0001	61,4881	0,0070	0,0001	102.3041
$\alpha_1$	-0,043	-0,0372	1,6597	-0,0224	-0,0264	1,2920	-0,0205	-0,0402	0,0475	-0.8450
$\alpha_2$	0,202	0,2039	0,0368	5,5436	0,2039	0,0369	5,5314	0,2034	0,0202	10.0836
$\alpha_3$	0,220	0,2223	0,0428	5,2000	0,2223	0,0431	5,1604	0,2200	0,0250	8.7918
$\rho_1$	-0,075	-0,1264	1,4524	-0,0870	-0,1223	1,1462	-0,1067	-0,0794	0,0366	-2.1693
$\rho_2$	0,000	-0,0056	0,0215	-0,2592	-0,0053	0,0220	-0,2432	-0,0010	0,0078	-0.1309
$\rho_3$	-0,126	-0,1256	0,0341	-3,6860	-0,1255	0,0356	-3,5253	-0,1255	0,0198	-6.3394
$\mu_1$	-0,005	-0,0071	0,0699	-0,1011	-0,0069	0,0549	-0,1251	-0,0052	0,0018	-2.8385
$\mu_2$	0,001	0,0011	0,0003	3,2101	0,0011	0,0003	3,2260	0,0010	0,0002	5.6038
$\mu_3$	0,005	0,0049	0,0019	2,5851	0,0049	0,0020	2,4709	0,0049	0,0011	4.5900
$\lambda$	0,040	0,0384	0,0060	6,4368	0,0385	0,0060	6,4058	0,0398	0,0019	21.4588
siml1		0,1577	0,1177		0,1587	0,1168		0,1647	0,0807	
siml2		0,5106	0,0684		0,5103	0,0676		0,5058	0,0478	
siml3		0,3308	0,0715		0,3301	0,0714		0,3291	0,0471	
estl1		0,1647	0,1240		0,1656	0,1231		0,1669	0,0824	
estl2		0,4803	0,1179		0,4806	0,1179		0,5003	0,0594	
estl3		0,3550	0,1288		0,3538	0,1290		0,3328	0,0579	

Tableau 5 ( $\rho_2 = 0, \lambda = 0,010$ )

Vraies Valeurs		[1] T=1000, R=1000			[2] T=2000, R=1000			[3] T=5000, R=1000		
		E()	Écart-type	Ratio t	E()	Écart-type	Ratio t	E()	Écart-type	Ratio t
$\sigma$	<b>0,007</b>	0,0070	0,0002	44,2873	0,0070	0,0001	62.2903	0,0070	0,0001	93.9129
$\alpha_1$	<b>-0,043</b>	-0,0362	0,0704	-0,5142	-0,0401	0,0699	-0.5734	-0,0435	0,0245	-1.7780
$\alpha_2$	<b>0,202</b>	0,0429	0,1252	0,3429	0,0389	0,1021	0.3809	0,0279	0,0772	0.3610
$\alpha_3$	<b>0,220</b>	0,0575	2,4609	0,0234	0,1521	0,5005	0.3039	0,1802	0,1898	0.9494
$\rho_1$	<b>-0,075</b>	-0,0906	0,0939	-0,9649	-0,0756	0,2307	-0.3276	-0,0764	0,0204	-3.7542
$\rho_2$	<b>0,000</b>	-0,0307	0,1181	-0,2601	-0,0057	0,0936	-0.0606	0,0176	0,0730	0.2410
$\rho_3$	<b>-0,126</b>	0,0911	5,2050	0,0175	-0,2209	3,2320	-0.0683	-0,1085	0,1957	-0.5542
$\mu_1$	<b>-0,005</b>	-0,0063	0,0071	-0.8820	-0,0055	0,0082	-0.6644	-0,0051	0,0019	-2.6787
$\mu_2$	<b>0,001</b>	-0,0036	0,0043	-0.8512	-0,0029	0,0033	-0.8835	-0,0021	0,0024	-0.8795
$\mu_3$	<b>0,005</b>	-0,0046	0,2381	-0.0192	0,0100	0,1514	0.0662	0,0044	0,0103	0.4247
$\lambda$	<b>0,010</b>	0,0525	0,0195	2.6877	0,0509	0,0182	2.7980	0,0472	0,0152	3.1083
siml1		0,7839	0,2553		0.8461	0,1791		0.9025	0,0967	
siml2		0,0152	0,0126		0.0103	0,0086		0.0067	0,0051	
siml3		0,1993	0,2459		0.1427	0,1722		0.0905	0,0926	
estl1		0,5560	0,2653		0.6277	0,2421		0.7301	0,2134	
estl2		0,3421	0,2643		0.2977	0,2551		0.2234	0,2270	
estl3		0,1019	0,1899		0.0746	0,1346		0.0465	0,0677	
Vraies Valeurs		[4] T=2000, R=1000(1%)			[5] T=2000, R=1000(NB)					
		E()	Écart-type	Ratio t	E()	Écart-type	Ratio t			
$\sigma$	<b>0,007</b>	0,0070	0,0001	62,5443	0,0070	0,0001	62.5914			
$\alpha_1$	<b>-0,043</b>	-0,0301	0,1067	-0,2818	0,0037	0,8017	0.0046			
$\alpha_2$	<b>0,202</b>	0,0991	0,2072	0,4782	0,2576	4,0264	0.0640			
$\alpha_3$	<b>0,220</b>	0,1481	0,2026	0,7313	0,1489	0,2069	0.7194			
$\rho_1$	<b>-0,075</b>	-0,1039	0,2792	-0,3723	-0,2684	4,9796	-0.0539			
$\rho_2$	<b>0,000</b>	-0,0205	0,2865	-0,0716	0,2638	10,9996	0.0240			
$\rho_3$	<b>-0,126</b>	-0,1012	0,3596	-0,2816	-0,0943	0,3096	-0.3046			
$\mu_1$	<b>-0,005</b>	-0,0085	0,0192	-0,4441	-0,0291	0,6160	-0.0472			
$\mu_2$	<b>0,001</b>	-0,0029	0,0057	-0,5109	-0,0023	0,0301	-0.0753			
$\mu_3$	<b>0,005</b>	0,0039	0,0177	0,2217	0,0036	0,0167	0.2172			
$\lambda$	<b>0,010</b>	0,0417	0,0294	1,4182	0,0401	0,0328	1.2229			
siml1		0,8461	0,1791		0,8461	0,1791				
siml2		0,0103	0,0086		0,0103	0,0086				
siml3		0,1427	0,1722		0,1427	0,1722				
estl1		0,6600	0,3170		0,6567	0,3307				
estl2		0,2414	0,3382		0,2429	0,3534				
estl3		0,0986	0,1586		0,1004	0,1593				

## 2. La prévision

Afin d'évaluer un modèle hors-échantillon il faut commencer par la division des observations en deux parties. La première partie, soit l'échantillon principal, sert à faire de l'inférence statistique sur les coefficients qui nous intéressent, alors que la deuxième partie, soit l'échantillon de comparaison, sert à comparer les prévisions hors-échantillon. Notons le nombre total d'observations par  $T$ , le nombre d'observations dans *l'échantillon principal* par  $R$  et l'horizon de prévision par  $h$ . Ainsi, le nombre d'observations dans *l'échantillon de comparaison* est  $T-R$ , tandis que le nombre de prévisions ponctuelles est  $T-R-h+1$ . Dans nos simulations, l'échantillon principal contient 90% de données et celui de comparaison contient le reste, soit 10% de données. Étant donné que nous nous concentrons ici sur un horizon d'une période ( $h=1$ ), le nombre de prévisions coïncide avec le nombre d'observations dans l'échantillon de comparaison, c.-à-d  $T-R$ . Donc, il y a  $T-R$  prévisions ponctuelles de  $y_{t+1}$ , que nous dénotons  $\hat{y}_{t+1}$ . De plus, la fenêtre glissante est la méthode de génération de prévisions ponctuelles à l'horizon d'une période. Alors, pour chaque prévision il y aura une erreur de prévision :

$$e_{t+1} = y_{t+1} - \hat{y}_{t+1},$$

où  $t$  est la dernière observation de l'échantillon principal.

À partir de l'erreur de prévision nous pouvons définir une fonction de perte  $L(e_{t+1})$ . Ici nous choisissons la fonction de perte quadratique :  $L(e_{t+1}) = e_{t+1}^2$ . Donc, si nous désirons comparer la performance hors-échantillon de deux modèles il faut comparer leurs fonctions de perte. Quand même, il faut toujours garder à l'esprit l'avertissement de Diebold (2012), dont nous avons parlé dans la revue de la littérature : l'objectif de base du test DM est d'établir si la différence entre deux séries des erreurs de prévisions est statistiquement significative, et l'implication que le modèle correspondant aux erreurs de prévision plus petites est meilleur que l'autre n'est pas nécessairement vraie. Notamment, le test n'était pas conçu pour établir si un modèle est meilleur que l'autre. Ensuite, il y aura deux séries des erreurs de prévisions (1 et 2) définies par la fonction de perte quadratique. Nous pouvons ainsi définir la différence  $d_{12t} = L(e_{1t}) - L(e_{2t})$  telle que  $\{d_{12t}\}_{t=1}^{T-R}$ . La question clé ici est de savoir si la différence entre la fonction de perte de deux séries des erreurs de prévisions est statistiquement significative. Le test Diebold et Mariano (1995) (DM) fournit une méthode pour répondre à cette question.

L'hypothèse centrale de ce test est que la série est faiblement stationnaire. Dès que cette hypothèse est satisfaite, la statistique de test est la moyenne échantillonnale divisée par l'écart-type estimé :



$$DM_{12} = \frac{\bar{d}_{12}}{\hat{\sigma}_{\bar{d}_{12}}} \rightarrow N(0,1),$$

où  $\bar{d}_{12} = \frac{1}{T-R} \sum_{i=1}^{T-R} d_{12i}$ . Puisque l'horizon de prévision ici est d'une période, nous prenons l'écart-type échantillonnaire. Nous obtenons les valeurs du numérateur et du dénominateur tout simplement en régressant le vecteur de différences sur la constante.

Comme la variable dépendante dans ce modèle SETAR à trois régimes est la différence, nous pouvons considérer celle-ci comme notre objet de prévision. Pourtant, pour comparer les prévisions nous prenons la variable en niveau. Pour chaque déplacement de la fenêtre glissante nous obtenons une prévision pour la période qui suit immédiatement la dernière période incluse. Ainsi, la prévision de la différence ( $\Delta \hat{y}_{t+1}$ ) est obtenue, où  $t$  est la dernière observation de la fenêtre glissante. Donc, la valeur prévue en niveau est donnée par  $\hat{y}_{t+1} = y_t + \Delta \hat{y}_{t+1}$ , où  $y_t$  est la dernière observation dans la fenêtre glissante. Dès que nous obtenons  $\{\hat{y}_{t+1}\}_{t=0}^{T-R}$ , nous pouvons définir la fonction de perte donnée plus haut. Ainsi, les erreurs de prévisions obtenues en estimant le modèle SETAR à trois régimes forment la série dénotée ci-haut par 2, c.-à-d.  $e_{2t}$ . Considérons aussi la série dénotée par 1 les erreurs de prévision d'une marche aléatoire ou un processus martingale. Par conséquent, nous définissons  $\hat{y}_{t+1} = y_t$  et  $e_{1t} = y_{t+1} - y_t$ . Ensuite, nous étudions la puissance du test DM quand le PGD est un SETAR à trois régimes.

Afin de tester la puissance du test DM nous considérons l'hypothèse unilatérale. Étant donné notre définition de la différence, une valeur positive signifie que la prévision obtenue à partir de SETAR à trois régimes, est « meilleure » que celle obtenue à partir d'une marche aléatoire. Alors, l'hypothèse nulle est  $H_0: DM_{12} \leq 0$  et celle alternative est  $H_1: DM_{12} > 0$ . Remarquez que  $DM_{12}$  est en fait le ratio  $t$  de la régression du vecteur de différences sur une constante quand la vraie valeur du paramètre est zéro.

Premièrement, nous considérons la puissance et le niveau (la taille) du test DM unilatéral pour les paramètres dans le tableau 3. Le modèle choisi sous l'hypothèse nulle est  $\Delta y_t = \sigma \varepsilon_t$ , qui est une marche aléatoire dans la forme la plus simple. Les résultats sont affichés dans le tableau 5.

Observez que la puissance est une fonction croissante de la taille de l'échantillon. C'est une conclusion logique de nos résultats sur l'estimation. Pour un échantillon plus grand, les estimateurs MCO sont plus forts. Autrement dit, la distribution de paramètres estimés obtenus est centrée très proche des vraies valeurs et son deuxième moment est plus petit. Évidemment, quand le PGD est un SETAR à trois régimes et la forme du modèle estimé est aussi un SETAR à trois régimes, nous obtenons des paramètres estimés plus proches de leurs vraies valeurs. Comme les paramètres estimés sont utilisés dans les prévisions ponctuelles, ces prévisions seront plus précises. En effet, vu que chacun des trois régimes a des

paramètres locaux, l'importance de la taille d'échantillon pour la qualité de la prévision doit être plus remarquable.

Il faut aussi observer que les tailles d'échantillon pour lesquelles les paramètres estimés sont proches des vraies valeurs (ex. 2000 ou 5000) ne sont pas typiques pour des séries macroéconomiques. Par conséquent, la puissance et le niveau du test plus pertinents dans le contexte macroéconomique se trouvent dans les colonnes correspondantes au  $T=1025$  dans les tableaux 5 et 6.

Tableau 5 : La puissance et le niveau empirique du test DM de 5% unilatéral

Descriptions	T=1025, R=300	T=1525, R=300	T=2025, R=300
<b>Puissance</b>	47,67%	65%	71,67%
<b>Niveau</b>	13%	12%	11,33%

#### i. La puissance à niveau corrigé (PNC)

Bien que nous obtenions la puissance du test DM, il faut remarquer que le niveau empirique est au-dessus de celui théorique (5%). L'implication intuitive ici est que quand le PGD est une marche aléatoire l'hypothèse nulle est sur-rejetée. Conséquemment, il y a un sur-rejet de l'hypothèse nulle aussi quand nous évaluons la puissance du test. Afin de régler ce sur-rejet de l'hypothèse nulle, il faut trouver une valeur critique qui corresponde au niveau de 5%. Dans ce cas nous nous attendons à ce que cette valeur soit plus élevée que celle d'une loi normale centrée réduite ( $\approx 1,6448$ ) pour un test unilatéral.

La valeur critique qui correspond au niveau de 5% est trouvée à l'aide de simulations pour lesquelles le PGD est la marche aléatoire. Nous procédons de la façon suivante : pour chaque répétition dans cette simulation nous obtenons une valeur de  $DM_{12}$ ; comme nous considérons plusieurs répétitions, nous obtenons finalement une liste de valeurs de  $DM_{12}$ ; la valeur critique au niveau de 5% est trouvée en identifiant la valeur de  $DM_{12}$  qui correspond au 95<sup>ème</sup> centile.

Tableau 6 : Puissance à niveau corrigé (PNC)

Descriptions	T=1025, R=300	T=1525, R=300	T=2025, R=300
<b>PNC</b>	28%	33%	55,33%
<b>Niveau</b>	5%	5%	5%
<b>Valeur critique (5%)</b>	2,10084	2,256434	2,05008

Le tableau 6 donne la puissance à niveau corrigé (PNC) pour les mêmes valeurs de la taille d'échantillon et du nombre de répétitions que dans le tableau 5. Comme nous avons attendu la PNC est plus basse que celle nominale.

Il faut bien remarquer que même quand le vrai modèle est un SETAR à trois régimes, et que nous l'estimons avec un SETAR à trois régimes, la supériorité de prévisions est attestée par le test DM

seulement dans un tiers des cas pour la taille d'échantillon inférieure à 1500. Donc, en se basant sur le test DM il est très facile d'en conclure que le modèle SETAR à trois régimes est inapproprié, même quand SETAR à trois régimes est le vraie modèle. Et même Diebold (2012) recommande l'évaluation dans l'échantillon plutôt que celle (pseudo-) hors-échantillon pour évaluer si un modèle décrit mieux ou moins bien les données.

Dans ce contexte nous comparons la performance du test de racine unitaire proposé par Bec, Ben Salem et Carrasco (2004) avec la performance du test DM. Nous imposons les mêmes restrictions sur les paramètres de SETAR à trois régimes sous l'hypothèse alternative. Ainsi, sous ces restrictions nous prenons  $\alpha_1 = \alpha_2 = \alpha_3 = \alpha$ ,  $\mu_1 = -\mu_3$  et  $\rho_1 = \rho_3$ , où  $\mu_1 \approx 1,3\lambda\rho_1$  en considérant toujours le modèle ci-dessous :

$$\Delta y_t = \begin{cases} \alpha_1 \Delta y_{t-1} + \mu_1 + \rho_1 y_{t-1} + \sigma \varepsilon_t, & y_{t-1} \leq -\lambda \\ \alpha_2 \Delta y_{t-1} + \mu_2 + \rho_2 y_{t-1} + \sigma \varepsilon_t, & |y_{t-1}| < \lambda, \\ \alpha_3 \Delta y_{t-1} + \mu_3 + \rho_3 y_{t-1} + \sigma \varepsilon_t, & y_{t-1} \geq \lambda \end{cases}$$

où  $\varepsilon_t$  est iid  $N(0,1)$ .

Nous commençons par trouver la valeur critique qui nous permettra de calculer la PNC. À cet effet, nous considérons le modèle  $\Delta y_t = 0,3\Delta y_{t-1} + \mu + 0,02\varepsilon_t$  comme PGD sous l'hypothèse nulle. Ensuite nous calculons les t-ratios (valeurs  $DM_{12}$ ) pour 10 000 répétitions pour la taille d'échantillon égale à 325. Nous trouvons que la valeur critique à 5% est 2,53023.

Les résultats sont donnés dans le tableau 7. Remarquez que les valeurs de paramètres dans le tableau sont normalisées. Nous voyons immédiatement que la PNC du test DM (DM\_PNC) est beaucoup plus faible que celle de supLR (supLR\_PNC) pour tous les cas considérés, soit les différentes valeurs de trois paramètres  $\alpha$ ,  $\rho_1$  et  $\lambda$ . Les 13 valeurs de supLR\_PNC proviennent de Bec, Ben Salem et Carrasco (2004). Nous observons pourtant que dans les cas de deux tests la puissance chute quand  $\rho_1$  s'approche de 0. De plus, dans les deux cas, la puissance est croissante en  $\alpha$  et  $\lambda$ .

Tableau 7 : La puissance nominale (DM\_P) et PNC (DM\_PNC) ( $\mu_2 = \rho_2 = 0$  et  $\sigma = 1$ )

	$\alpha$	$\rho_1$	$\lambda$	DM_P	DM_PNC	supLR_PNC
1	0	-0.3	10	26.63	4.73	82.8
2	0	-0.3	5	18.67	2.53	70.4
3	0	-0.3	2	18.43	2.47	80
4	0	-0.1	10	14.07	1.7	17.1
5	0	-0.1	5	14.93	1.5	15.7
6	0	-0.1	2	15.47	1.7	21.4
7	0	-0.05	10	13.27	1.6	8.3
8	0	-0.05	5	11.87	1.43	9.2
9	0	-0.05	2	13.67	1.73	10.9
10	0.3	-0.3	10	33.67	7.63	94.1
11	0.3	-0.3	5	29.6	6.33	87.8
12	0.3	-0.3	2	30.43	6.3	
13	0.3	-0.1	10	27.73	5.83	25.5
14	0.3	-0.1	5	25.77	5.5	25.8
15	0.3	-0.1	2	26.8	5.43	
16	0.3	-0.05	10	25.93	5.37	
17	0.3	-0.05	5	26.27	5.4	
18	0.3	-0.05	2	25.43	5.5	

En comparant les deux séries de prévisions provenant de deux modèles différents nous aurions pu implicitement tirer des conclusions sur la qualité des deux modèles. Prenons le cas des prévisions d'un modèle SETAR à trois régimes et celles d'une marche aléatoire. En extrapolant les conclusions d'une comparaison de prévisions, soit une analyse (pseudo-)hors-échantillon, sur la capacité de chacun de modèles à décrire les données, nous aurions pu établir que la marche aléatoire décrit mieux les données qu'un modèle SETAR à trois régimes. Par extension, nous aurions pu en tirer la conclusion que le processus stochastique a une racine unitaire.

Quand même, en regardant les résultats dans le tableau 7, particulièrement la colonne contenant la PNC du test DM (DM\_PNC), il est évident que ce test n'est pas le meilleur choix pour ce genre de conclusion ou pour établir la supériorité d'un modèle. Notamment, le test supLR proposé par Bec, Ben Salem et Carrasco (2004) a une PNC beaucoup plus forte. C'est surtout le cas dans les échantillons assez petits (250-600 observations), une taille habituelle pour les séries temporelles macroéconomiques. Finalement, ces résultats renforcent les l'avertissement préconisés par Diebold (2012).

Dans ce contexte nous nous attendons qu'en appliquant le test DM aux prévisions basées sur les taux de change réels pour lesquels Bec, Ben Salem et Carrasco (2004) ont rejeté l'hypothèse nulle (il y a une RU), nous ne serons pas capables de rejeter l'hypothèse nulle d'un test DM bilatéral. À cet effet, nous comparons les prévisions obtenues à partir d'un modèle SETAR à trois régimes et celles obtenues à partir de la marche aléatoire. Dans le tableau 8 nous présentons les résultats d'application du test DM bilatéral dans le contexte d'une analyse (pseudo-) hors-échantillon pour les taux de change réels pour lesquels Bec, Ben Salem et Carrasco (2004) ont rejeté l'hypothèse nulle. Nous considérons deux tailles de *l'échantillon de comparaison* : 10% et 20% d'observations. Dans les deux cas il y a respectivement 290 et 260 observations dans *l'échantillon principal* et 35 et 65 observations dans *l'échantillon de comparaison*.

Tableau 8 : Le test DM appliqué aux taux de change réels

Taux de change réel (/DEM)	p - valeur	
	10%	20%
<b>FRF</b>	0,248	0,023
<b>ITL</b>	0,025	0,91
<b>BEF</b>	0,357	0,143
<b>NLG</b>	0,512	0,021*
<b>ESP</b>	0,126	0,004

Observez que les p-valeurs pour quatre taux de change sont au-dessous de 5%. Cependant, il faut remarquer que l'hypothèse nulle  $H_0: DM_{12} = 0$  est rejetée en faveur des prévisions provenant du modèle SETAR à trois régimes seulement dans le cas de florin hollandais (NLG). Ceci n'est pas surprenant vu la puissance nominale et la PNC du test DM que nous avons trouvé. Donc, il serait inapproprié d'en tirer des conclusions défavorables sur la capacité du modèle SETAR à trois régimes à décrire la dynamique de taux de change réel en question.

## V. CONCLUSION

Dans cette étude nous trouvons que pour la taille d'échantillon habituelle en macroéconomie, soit 325 observations, le test DM a une puissance à niveau corrigé (PNC) très faible quand le PGD est un SETAR à trois régimes. Par conséquent, si nous comparons à l'aide du test DM les prévisions provenant d'un modèle SETAR à trois régimes et celles provenant d'une marche aléatoire, et le résultat suggère que les prévisions obtenues à partir de la marche aléatoire sont meilleures, c'est incorrect d'en tirer la conclusion que le modèle SETAR à trois régimes ne décrit pas bien la dynamique de la variable visée.

A cet effet, l'inférence en échantillon est une meilleure approche que la prévision hors-échantillon. Pour le modèle SETAR à trois régimes nous considérons le test de racine unitaire (supLR) proposé par Bec, Ben Salem et Carrasco (2004). La puissance de ce test est beaucoup plus élevée que la puissance de test DM. Et donc, les résultats de test supLR sont plus fiables quand nous désirons d'établir si le modèle SETAR à trois régimes est un bon choix. Il faut ajouter qu'il est possible que pour des échantillons plus grands la différence entre la performance de deux tests soit moins importante. Cette affirmation est fondée sur l'analyse, présentée au début de la section « La prévision », de la relation entre la taille d'échantillon et la qualité de prévision.

Finalement, l'étude sur la puissance nominale et la PNC du test DM peut être appliquée aux autres modèles afin d'établir s'il y a en effet des cas dans lesquels la comparaison de prévisions à l'aide du test DM reflète la qualité de modèles. Autrement dit, nous n'excluons pas la possibilité que pour certains modèles la supériorité de prévisions, établie à l'aide du test DM, implique la supériorité de modèle. Pourtant, nous considérons cette possibilité peu plausible. De plus, il est aussi possible de considérer des prévisions à un plus grand horizon que celui utilisé ici, soit une période.

## Références bibliographiques:

- Bec, F., M. Ben Salem, et M. Carrasco. 2004. «Tests for unit-root versus threshold specification with an application to the purchasing power parity relationship.» *Journal of Business & Economic Statistics* 22 (4): 382-95.
- Chen, Jian, et Nelson C. Mark. 1996. «Alternative Long-horizon Exchange-rate Predictors.» *International Journal of Finance & Economics* 1 (4): 229-50.
- Clark, T. E., et K. D. West. 2007. «Approximately normal tests for equal predictive accuracy in nested models.» *Journal of Econometrics* 138 (1): 291-311.
- Davies, R. B. 1987. «Hypothesis-Testing When a Nuisance Parameter Is Present Only under the Alternative.» *Biometrika* 74 (1): 33-43.
- Diebold, F. X., et R. S. Mariano. 1995. «COMPARING PREDICTIVE ACCURACY.» *Journal of Business & Economic Statistics* 13 (3): 253-63.
- Diebold, F.X. 2012. «Comparing Predictive Accuracy, Twenty Years Later: A personal perspective on the Use and Abuse of Diebold-Mariano Tests.». Department of Economics, University of Pennsylvania.
- Hansen, B. E. 1996. «Inference when a nuisance parameter is not identified under the null hypothesis.» *Econometrica* 64 (2): 413-30.
- Mark, N. C. 1995. «EXCHANGE-RATES AND FUNDAMENTALS - EVIDENCE ON LONG-HORIZON PREDICTABILITY.» *American Economic Review* 85 (1): 201-18.
- Meese, R. A., et K. Rogoff. 1983. «EMPIRICAL EXCHANGE-RATE MODELS OF THE SEVENTIES - DO THEY FIT OUT OF SAMPLE.» *Journal of International Economics* 14 (1-2): 3-24.
- Molodtsova, T., et D. H. Papell. 2009. «Out-of-sample exchange rate predictability with Taylor rule fundamentals.» *Journal of International Economics* 77 (2): 167-80.